PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H01J 49/04, C12Q 1/00, G01N 33/48 — A1

(11) International Publication Number: **WO 96/36986**

(43) International Publication Date: 21 November 1996 (21.11.96)

(21) International Application Number: PCT/US96/07146

(22) International Filing Date: 17 May 1996 (17.05.96)

(30) Priority Data:

08/446,055	19 May 1995 (19.05.95)	US
08/447,175	19 May 1995 (19.05.95)	US

(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant: PERSEPTIVE BIOSYSTEMS, INC. [US/US]; 500 Old Connecticut Path, Framingham, MA 01701 (US).

(72) Inventors: PATTERSON, Dale, H.; Apartment J, 38 Bay Ridge Drive, Nashua, NH 03062 (US). TARR, George, E.; 640 Essex Street, S. Hamilton, MA 01982 (US).

(74) Agent: TURANO, Thomas, A.; Testa, Hurwitz & Thibeault, L.L.P., High Street Tower, 125 High Street, Boston, MA 02110 (US).

(54) Title: METHODS AND APPARATUS FOR SEQUENCING POLYMERS WITH A STATISTICAL CERTAINTY USING MASS SPECTROMETRY

(57) Abstract

The method and apparatus disclosed herein are useful for sequencing polymers using mass spectrometry. The methods involve differing ratios of hydrolyzing agent to polymer disposed upon a reaction surface adapted for use with a mass spectrometer. The methods further involve integrating data obtained from mass spectrometry analysis of a plurality of series of hydrolyzed polymer fragments, and provide statistical interpretation paradigms and computer software therefor. The apparatus involves a mass spectrometer sample holder, having hydrolyzing agent disposed thereon, which is useful for adapting any mass spectrometer for polymer sequencing.

Table 1

Peptide	SEQ ID Nos.	Sequence	Average Mass[1]	Charge[2]	Polarity
Sleep Inducing Peptide	1	WAGGDASGE	848.8	-2.0	polar
Amino Terminal Region of Hbs β chain[3]	2	VHLTPVEK	922.1	+0.5	mid
Interleukin-1β 163-171 Fragment[3]	3	VQGEESNDK	1005.0	-2.0	polar
TRH Precursor	4	KRQHPGKR	1006.2	+4.5	very
Bradykinin	5	RPPGFSPFR	1061.2	+2.0	mid
Lutenizing Hormone Releasing Hormone[3]	6	pyro.EHWSYGLRPG.amide	1182.3	+1.5	mid
Physalaemin	7	pyro.EADPNKFYGLM.amide	1265.4	0	mid
Angiotensin 1	8	DRVYIHPFHL	1295.5	+1.0	non
Renin Inhibitor	9	PHPFHFFVYK	1318.5	+2.0	non
Kassinin	10	DVPKSDQFVGLM.amide	1334.5	-2.0	non
Substance P	11	RPKPQQFFGLM.amide	1347.6	+3.0	mid
T-Antigen Homolog	12	CGYGPKKKRKVGG	1377.7	+5.0	polar
Osteocalcin 7-19 Fragment	13	GAPVPYPDPLEPR	1407.6	-1.0	mid
Fibrinopeptide A	14	ADSGEGDFLAEGGGVR	1536.6	-3.0	mid
Thymopoietin II 29-41 Fragment	15	GEQRKDVYVQLYL	1610.8	0	polar
Bombesin	16	pyro.EQRLGNQW(AVGH)LM.amide	1619.9	+1.5	mid
ACTH 11-24 Fragment	17	KPVGKKRRPVKVYP	1652.1	+6.0	mid
α-Melanocyte Stimulating Hormone	18	acetyl.STSMEHFRWGKPV. amide	1664.9	+1.5	mid
Angiotensinogen 1-14 Fragment	19	DRVYIHPFHLLVYS	1759.0	+1.0	non
Angiogenin	20	ENGLPVHLDQSI(FR)R	1781.0	+0.5	mid
Glucagon	21	HSQ...DSRRAQDFVQW(LMN)T	3482.8	+1.0	polar
ACTH7-38 Fragment	22	FRW...RRPVKVYPNGAEDESAEAF PLE	3659.15	+2.0	polar

1 calculated

2 at pH 6.5

3 no sequence information was obtained

Listed in Table 1 are the peptides that have been digested and analyzed using this novel on-plate strategy. These peptides were selected to represent peptides of varying amino acid composition, size (up to MW = 3659.15), charge and polarity. The bolded amino acids indicate that a peak representing the loss of that residue was observed in one or more of the MALDI spectra taken across the row of digestions. In order to be able to identify a residue, the peak representing the loss of that amino acid and the preceding amino acid must be present. The residues that are enclosed in parenthesis are those for which the sequence order could not be deduced. Overall, CPY offered some sequence information from the C-terminus for most of the peptides digested, lending no sequence information in only three of the 22 cases. In two of these three cases, the C-terminus was a lysine followed by an acidic residue at the penultimate position. CPY has been reported to possess reduced activity towards basic residues at the C-terminus, and the presence of the neighboring acidic residue seems to further reduce its activity. In the case of the lutenizing hormone releasing hormone (LH-RH), the C-terminal amidated glycine followed by proline at the penultimate position inhibited CPY activity which agrees with reports of CPY slowing at both proline and glycine residues (Hayashi et al. (1975) J. Biochem. 77:69-79;

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Arg Pro Lys Pro Gln Gln Phe Phe Gly Leu Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Cys Gly Tyr Gly Pro Lys Lys Lys Arg Lys Val Gly Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Ala Pro Val Pro Tyr Pro Asp Pro Leu Glu Pro Arg
1 5 10

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ala Asp Ser Gly Glu Gly Asp Phe Leu Ala Glu Gly Gly Gly Val Arg
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gly Glu Gln Arg Lys Asp Val Tyr Val Gln Leu Tyr Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Gln Arg Leu Gly Asn Gln Trp Ala Val Gly His Leu Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Lys Pro Val Gly Lys Lys Arg Arg Pro Val Lys Val Tyr Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids